

Fig. 1: Amino acid sequence of the HBsAg a determinant of the different HBV genotypes as compared with the novel variant HDB 11

A representative genome was used as the basis for each genotype and the aa sequence was deduced from the nucleotide sequence

A: X70 185; B: D00331; C: X01587; D: X72702, E: X75664; F: X75663; G: FRI
(Stuyver et al.; J. Gen. Virol. 81: 67-74 (2000); Norder et al.; J. Gen. Virol. 73: 3141-3145 (1992))

aa #	101	111	121	131	141	151	161	170
Genotype								
A	QGMLPVCPLI	PGSTTTSTGP	CKTCTTPAQG	NSMFPSCCCT	KPTDGNCTCI	PIPSSWAFK	YLWEWASVRF	
B	-----	---S-----	-----	T-----	-----	-----	-----	
C	-----	L--TS-----	-----	T-----	-----	-----	R F--G-----	
D	-----	---S-----	---R-----	T--Y-----	---S-----	---G--F-----	A-----	
E	-----	---S-----	---R--M-L---	T-----	---S-----	---G--F-----	A-----	
F	-----	-----	---L---	T-----	---S-----	---LG-	A-----	

HDB 11 --I----- --AINNR-Q-K-----T-H- T--Y-Y-----S-----G- F-----A--
(R)

aa # 103 114 120 129 136 143 159 168

The amino acid substitutions which differ from the wild-type genotype D, ayw2, HBV are printed in bold type

BEST AVAILABLE COPY

Fig. 2 Nucleotide sequence of the S gene of the known HBV ayw2 wild type

encoding the HBV surface protein (surface antigen, HBsAg), and resulting amino acid sequence in the 3-letter and, especially, 1-letter codes (Coleman et al; WO 02/079217 A1) Continuous numbering of nucleotides (nt) encoding the surface antigen (excl. pre S1 and pre S2 regions) Continuous numbering of amino acids (aa)

	(nt)	(aa)
1	60	20
61	120	40
121	180	60
181	240	80
241	300	100
301	360	120
361	420	140
421	480	160
481	540	180
541	600	200
601	660	220
661		
231		

ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG TTA CAG GCG GGG TTT TTC
Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln Ala Gly Phe Phe
M E N I T S G F L G P L L V L Q A G F F
TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT
Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser Thr Thr Ser Leu Asn
L L T R I L T I P Q S L D S W W T S L N
TTT CTA GGG GGA ACT ACC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC
Phe Leu Gly Thr Thr Val Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His
F L G O T T V C L G Q N S Q S P T S N H
TCA CCA ACC TCC TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT
Ser Pro Thr Ser Cys Pro Pro Thr Cys Pro Gly Tyr Arg Thr Met Cys Leu Arg Arg Phe
S P T S C P T C P Q Y R W M C L R R F
ATC ATC TTC CTC TTC ATC CTC CTG CTA TGC CTC ATC TTC TTG TTG GTT CTT CTG GAC TAT
Ile Ile Phe Leu Phe Ile Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr
I I F L F I L L L C L I F L L V L L D Y
CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA TCA ACC ACC ACC GGA CCC
Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Thr Thr Thr Gly Pro
Q G M L P V C P L I P G S T T S T G P
TGC AGA ACC TGC ACG ACT CCT GCT CAA GGA ACC TCT ATG TAT CCC TCC TGT TGC TGT ACA
Cys Arg Thr Cys Thr Thr Pro Ala Gln Gly Thr Ser Met Tyr Pro Ser Cys Cys Thr
C R T C T P A Q G T S M Y P S C C T
AAA CCT TCG GAT GGA AAC TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA
Lys Pro Ser Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Thr Ala Phe Gly Lys
K P S D G N C T C I P I S S W A F G K
TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT
Phe Leu Thr Glu Thr Ala Ser Ala Arg Phe Ser Thr Leu Ser Leu Leu Val Phe Val
F L W B W A S A R F S W L S L V P F V
CAO TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG ATG TGG TAT
Gln Thr Phe Val Gly Leu Ser Pro Thr Val Thr Leu Ser Val Ile Thr Met Thr Tyr
Q W F V G L S P T V W L S V I W M W Y
TGG GGG CCA AGT CTG TAC TCC ATC TTG AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT
Trp Gly Pro Ser Leu Tyr Ser Ile Leu Ser Pro Phe Leu Pro Leu Leu Phe Phe
W G P S L Y S I L S P P L L L L P I F P
TGT CTT TGG GTA TAC ATT 678
Cys Leu Thr Val Tyr Ile 226
C L W V Y I

BEST AVAILABLE COPY

Fig. 3 Nucleotide sequence of the HBV surface antigen-encoding S gene of the HBV ayw2 wild type (upper row of nt 1 to nt 678) as compared with the nucleotide sequence, which is sequenced from nt 127 to nt 588, of the novel variant HDB 11 (lower row, in which nucleotide differences are printed in bold type and the mutations which do not lead to any amino acid substitution are bracketed)

1	ATG GAG AAC ATC ACA TCA GGA TTC CTA GGA CCC CTG CTC GTG ITA CAG GCG GGG TTT TTC	60
61	TTG TTG ACA AGA ATC CTC ACA ATA CCG CAG AGT CTA GAC TCG TGG TGG ACT TCT CTC AAT	120
121	TTT CTA GGG GGA ACT ACC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC	180
181	- 127: GGG GGA ACT ACC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC	240
241	TCA CCA ACC TCC TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT	300
301	TCA CCA ACC TCC TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT	360
361	ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GTT CTT CTG GAC TAT	420
421	ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GCT CTT CTG GAC TAT	480
481	CAA GGT ATG TTG CCC GTT TGT CCT CTA ATT CCA GGA TCA TCA ACC ACC AGC ACG GGA CCC	540
541	CAA GGT ATA TTG CCC GTT TGT CCT CTA ATT CCA GGA(TCT)GCA ATC AAC AAC AGG GGACAA	600
601	TGC AAA ACC TGC ACG ACT CCT GCT CAA GGA ACC TCT ATG TAT CCC TCC TGT TGC TGT ACA	660
661	TGC AAA ACC TGC ACG ACT ACT GCT CAC GGA ACC TCT ATG TAT CCC TAC TGT TGC TGT (ACC)	
	AAA CCT TCG GAT GGA AAC TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA	
	AAA CCT TCG(GAC)GGA(AAT)TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA	
	TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC TCT TGG CTC AGT TTA CTA GTG CCA TTT GTT	
	TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC(TCC)TGG CTC AGT TTA CTA GTT(CCC)TTT GTT	
	CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG ATG ATG TGG TAT	
	CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG 588	
	TGG GGG CCA AGT CTG TAC TCC ATC TTG AGT CCC TTT TTA CCG CTG TTA CCA ATT TTC TTT	
	TGT CTT TGG GTA TAC ATT 678	

Fig. 4 Nucleotide sequence of the S gene of the novel HBV variant HDB 11:
nt 127 to nt 588 of the HBV surface antigen-encoding genome.
Only the nucleotide differences which lead to a change in the amino
acid sequence are printed in bold.

```
- 127:  GGG GGA ACT ACC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC 180
181  TCA CCA ACC TCC TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT 240
241  ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GCT CTT CTG GAC TAT 300
301  CAA GGT ATA TTG CCC GTT TGT CCT CTA ATT CCA GGA TCT GCA ATC AAC AAC AGG GGACAA 360
361  TGC AAA ACC TGC ACG ACT ACT GCT CAC GGA ACC TCT ATG TAT CCC TAC TGT TGT TGT ACC 420
      (AGA, 364-366)
421  AAA CCT TCG GAC GGA AAT TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA 480
481  TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC TCC TGG CTC AGT TTA CTA GTT CCC TTT GTT 540
541  CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG 588
```

Fig. 5 S gene nucleotide sequence (nt 127 to 588) and corresponding amino acid sequence (aa 43 to 196) of the novel HBV variant HDB 11 (amino acids which are substituted as compared with the HBV ayw2 wild type are printed in bold and underlined)

```

- 127: GGG GGA ACT ACC GTG TGT CTT GGC CAA AAT TCG CAG TCC CCA ACC TCC AAT CAC      180
- aa 43: G G T T V C L G Q N S Q S P T S N H      60

181      TCA CCA ACC TCC TGT CCT CCA ACT TGT CCT GGT TAT CGC TGG ATG TGT CTG CGG CGT TTT      240
61      S P T S C P P T C P G Y R W M C L R R F      80

241      ATC ATC TTC CTC TTC ATC CTG CTG CTA TGC CTC ATC TTC TTG TTG GCT CTT CTG GAC TAT      300
81      I I F L F I L L C L I F L L A L L D Y      100

301      CAA GGT ATA TTG CCC GTT TGT CCT CTA ATT CCA GGATCT GCA ATC AAC AAG GGACAA      360
101      Q G I L P V C P L I P G S A I N N R G Q      120

361      TGC AAA ACC TGC ACG ACT ACT GCT CAC GGA ACC TCT ATG TAT CCC TAC TGT TGC TGT ACC      420
121      C K T C T T I A H G T S M Y P Y C C C T      140
      (AGA)
      (R in a 122)

421      AAA CCT TCG GAC GGA AAT TGC ACC TGT ATT CCC ATC CCA TCA TCC TGG GCT TTC GGA AAA      480
141      K P S D G N C T C I P I P S S W A F G K      160

481      TTC CTA TGG GAG TGG GCC TCA GCC CGT TTC TCC TGG CTC AGT TTA CTA GTT CCC TTT GTT      540
161      F L W E W A S A R F S W L S L L V P F V      180

541      CAG TGG TTC GTA GGG CTT TCC CCC ACT GTT TGG CTT TCA GTT ATA TGG      588
181      Q W F V G L S P T V W L S V I W -196

```

The following aa are substituted (x) in the HDB 11 variant as compared with the HBV ayw2 wild type:
V 96 (A) (not in the region of the a determinant),
M 103 (I), S 114 (A), T 115 (I), T 116 (N), S 117 (N), T 118 (R), P 120 (Q), T 127 (T), Q 129 (H) and
S 136 (Y) (all in the region of the a determinant)

BEST AVAILABLE COPY

Abb.6 Vergleich der Aminosäure-Sequenzen der a-Determinante (aa 100 bis aa 180) der neuen Variante HDB 11 (untere Reihe) mit dem Wildtyp HBV ayw2 (obere Reihe)

	aa-Sequenz Wildtyp ayw2 : aa-Sequenz Variante HDB 11 :																		100
	Y	Y																	
101	Q	G	M	L	P	V	C	P	L	I	P	G	S	S	T	T	S	T	G
	Q	G	I	L	P	V	C	P	L	I	P	G	S	A	I	N	N	R	P
																			Q
121	C	R	T	C	T	T	P	A	Q	G	T	S	M	Y	P	S	C	C	T
	C	K(R)	T	C	T	T	T	A	H	G	T	S	M	Y	P	Y	C	C	T
141	K	P	S	D	G	N	C	T	C	I	P	I	P	S	S	W	A	F	G
	K	P	S	D	G	N	C	T	C	I	P	I	P	S	S	W	A	F	G
161	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	V
	F	L	W	E	W	A	S	A	R	F	S	W	L	S	L	L	V	P	V

Die folgenden aa sind gegenüber dem Wildtyp HBV ayw2 bei der HDB 11-Variante substituiert (x):
 V 96 (A) (nicht in der Region der a-Determinante),
 M 103 (I), S 114 (A), T 115 (I), T 116 (N), S 117 (N), T 118 (R), P 120 (Q), P 127 (T), Q 129 (H) und
 S 136 (Y) (alle in der Region der a-Determinante)

BEST AVAILABLE COPY
 BEST AVAILABLE COPY